DOCUMENT RESUME

ED 115 760 CE 005 580

AUTHOR Phalen, William J.
TITLE Comprehensive Occupational Data Analysis Programs (CODAP): Ordering of Hierarchically Grouped Case Data (KPATH) and Print KPATH (PRKPTH) Programs.
INSTITUTION Air Force Human Resources Lab., Lackland AFB, Tex. Occupational and Manpower Research Div.
SPONS AGENCY Air Force Human Resources Lab., Brooks AFB, Texas.
PUB DATE Aug 75
NOTE 30p.

EDRS PRICE MF-$0.76 HC-$1.95 Plus Postage
DESCRIPTORS Computer Oriented Programs; *Computer Programs; *Data Analysis; Data Collection; *Data Processing; Information Processing; Information Retrieval; *Job Analysis; *Occupational Information; Statistical Data
IDENTIFIERS CODAP; *Comprehensive Occupational Data Analysis Programs

ABSTRACT

The Comprehensive Occupational Data Analysis Programs (CODAP) are designed for analyzing and retrieving occupational survey information collected by means of job inventories. Two programs that are parts of the CODAP system are described: the Ordering of Hierarchically Grouped Case Data (KPATH) program which reorders case data permitting identification of background variables having similar values for all or most cases in the group and enabling the generation of composite job descriptions for any hierarchical group; and the Print KPATH (PRKPTH) program which produces a printed report of case data elements for any selected background variables from an input-ordered or KPATH-ordered history data tape. The report discusses the application of these programs with supporting diagrams, tables, charts, and examples. Appended materials consisting of descriptions of other CODAP programs and a definition of terms used in the KPATH program documentation are included. (Author/EC)

AFHRL-TR-75-32

CE

AIR FORCE HUMAN RESOURCES LABORATORY

COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS (CODAP): ORDERING OF HIERARCHICALLY GROUPED CASE DATA (KPATH) AND PRINT KPATH (PRKPTH) PROGRAMS

By
William J. Phalen

OCCUPATIONAL AND MANPOWER RESEARCH DIVISION
Lackland Air Force Base, Texas 78236

August 1975
Interim Report for the Period 2 March 1973 – 1 February 1975

Approved for public release; distribution unlimited.

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This interim report was submitted by Occupational and Manpower Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236, under project 7734, with the Hq Air Force Human Resources Laboratory, Brooks Air Force Base, Texas 78235.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved.

JAMES B. CARPENTER, Colonel, USAF
Chief, Occupational and Manpower Research Division

Approved for publication.

HAROLD E. FISCHER, Colonel, USAF
Commander

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFHRL-TR-75-32
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS (CODAP). ORDERING OF HIERARCHICALLY GROUPED CASE DATA (KPATH) AND PRINT KPATH (PRKPTH) PROGRAMS
5. TYPE OF REPORT & PERIOD COVERED: Interim, 2 March 1973 – 1 February 1975
6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): William J. Phalen
8. CONTRACT OR GRANT NUMBER(s)

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Occupational and Manpower Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62703F, 77340116

11. CONTROLLING OFFICE NAME AND ADDRESS: Hq Air Force Human Resources Laboratory (AFSC), Brooks Air Force Base, Texas 78235
12. REPORT DATE: August 1975
13. NUMBER OF PAGES: 30

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

clustering procedures
CODAP programs
grouping procedures
hierarchical grouping
job analysis
job survey
KPATH program
occupational analysis
occupational research
occupational survey
PRKPTH program
task analysis
work analysis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This is one in a series of reports written to acquaint occupational analysts, occupational research personnel, and personnel managers with the functions and utilities of a set of Comprehensive Occupational Data Analysis Programs (CODAP). This technical report describes and gives examples of applications of the Ordering of Hierarchically Grouped Case Data (KPATH) and Print KPATH (PRKPTH) programs. These programs greatly facilitate the identification of background variables having similar values for all or most cases in any hierarchical group. Because of the highly technical nature of this report, it will be of interest only to readers who have a need to understand how the CODAP KPATH and PRKPTH programs function.

PREFACE

Over the past twelve years, the Air Force Human Resources Laboratory's Occupational Research and Manpower Division has developed a Comprehensive set of Occupational Data Analysis Programs (CODAP). These programs are designed for analyzing and retrieving occupational survey information collected by means of job inventories. The CODAP system was originally written for execution on an IBM 7040 computer. Many of the CODAP routines were coded in machine language (MAP), thus making it difficult for agencies not having access to this specific piece of equipment to make use of these powerful analysis programs. The United States Marine Corps had the CODAP system rewritten for execution on an IBM 360-65 computer for application by their service; and the Office of the Assistant Secretary of Defense for Manpower and Reserve Affairs had a version written for an IBM 370-155 computer for use by all the military services, as well as the United States Coast Guard. The United States Navy was designated as the executive agent for this version of CODAP. As the CODAP system becomes available to agencies outside the Air Force, there is a recognized need for the specific CODAP programs to be described in some detail.

It is not feasible to describe the entire system in a single document. This technical report describes the Ordering of Hierarchically Grouped Case Data (KPATH) and Print KPATH (PRKPTH) programs. The concept of KPATH ordering of hierarchically grouped data was developed by Dr Joe H. Ward, Jr., who received programming assistance from Mr Wayne E. Fisher and Mr Daniel Rigney. Dr Raymond E. Christal was responsible for the application of hierarchical grouping to the analysis of occupational data within the CODAP system. Programming of the CODAP KPATH and PRKPTH programs was accomplished under contract with the Computer Sciences Corporation.

This study was conducted under Project 7734, Development of Methods for Describing, Evaluating, and Structuring Air Force Occupations; Task 77340116, Development and Modification of Comprehensive Occupational Data Analysis Programs (CODAP).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES *(Continued)*

COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS (CODAP): ORDERING OF HIERARCHICALLY GROUPED CASE DATA (KPATH) AND PRINT KPATH (PRKPTH) PROGRAMS

I. INTRODUCTION

When cases have been grouped hierarchically on similarity of work performed, as reported in a job inventory, the background and task data for each case in any one of the hierarchical groups can be placed in proximity to that of other cases in the group by the Ordering of Hierarchically Grouped Case Data (KPATH)[1] program. Such a reordering of case data permits identification of background variables having similar values for all or most cases in the group, and it enables the generation of composite job descriptions for any hierarchical group, since the case data are in a readily accessible sequence.

While understanding of the overall purpose of the KPATH program is not particularly difficult, the mechanics of the program and its applications as an analytical tool are not so easily visualized without some degree of familiarity with distinct CODAP programs which help produce the input to KPATH or which use its output. Since previous technical reports deal with these associated programs, they will not be discussed in detail However, the relationship of each of the programs to KPATH is stated briefly, and reference is made to the report which discusses the program in depth.

Description of the KPATH program will occupy the first part of this report; the second topic discussed will be the Print KPATH (PRKPTH) program, which produces a printed report of case data elements for any selected background variables from an input-ordered or KPATH-ordered history data tape. In addition, a brief description of each program is provided in Appendix A for the convenience of the reader Appendix B is intended as a handy glossary of terms which describe the functional elements of the KPATH program; e.g., "hierarchical grouping".

II. DESCRIPTION AND APPLICATION OF THE PROGRAMS

KPATH Program

The purpose of the KPATH program is to sequence occupational survey case data in a so-called "KPATH" order. KPATH ordering is accomplished by sequencing cases in such a way that those cases or groups of cases which merged at each stage of the hierarchical clustering process are positioned adjacently. The sequential numbers assigned to cases so positioned are referred to as "KPATH sequence numbers." Perhaps, the concept of "KPATH order" and "KPATH sequence number" can best be illustrated by an example.

Consider a sample of 12 cases with manually assigned case control numbers of 01 through 12 (called "external" case control numbers). The punched data cards for those cases are placed in case number sequence and in card number sequence within cases by a program called Setcheck (SETCHK) which also eliminates cases without the proper number and sequence of cards per case. The punched cards for the surviving cases are then read into a checking and formatting program called Input Standard (INPSTD), which is described in Christal (1972). All cases which survive the checking portion of the INPSTD routine are assigned consecutive sequence numbers by the computer corresponding to the case input order (called "internal" case control numbers). Thus, if cases 03, 06, and 11 were rejected in INPSTD, the computer-assigned numerical identifiers (ID's) for the remaining nine cases would be. 01=1, 02=2, 04=3, 05=4, 07=5, 08=6, 09=7, 10=8, and 12=9. The nine acceptable cases are fed into an overlap program (OVRLAP) which computes a job similarity index for all possible pairs of the nine cases.

Following the computation of the overlap matrix, a clustering program (GROUP) is applied which combines cases and groups of cases according to their similarity index by stages until all cases have been merged into a single group. A thorough explanation of the mathematical processes used in the OVRLAP and GROUP programs is given in Archer (1966)[2]. It is in the GROUP program that sequencing techniques are applied that permit the assignment of an appropriate KPATH sequence number to each case. How the computer-assigned case numbers are manipulated during the clustering process to generate the appropriate KPATH sequence number for each case will now be discussed.

Keeping in mind that each individual case is initially a single-member "group," and that the computer-assigned case number is also the "group" ID, the following rule for sequencing groups is applied at

each stage of the clustering process. If a lower-numbered group (IBEST) is merged with a higher-numbered group (JBEST), the lower-numbered group is placed first in sequence, and the newly formed group is identified by the ID of the lower-numbered group. This rule is illustrated by Figure 1, in which the nine cases are grouped on the basis of a previously generated similarity matrix. This matrix is not shown, but you may refer to Appendix B for the definition of a similarity matrix as used in CODAP and to Archer (1966) for a thorough explanation of its computation in OVRLAP and its use in GROUP.

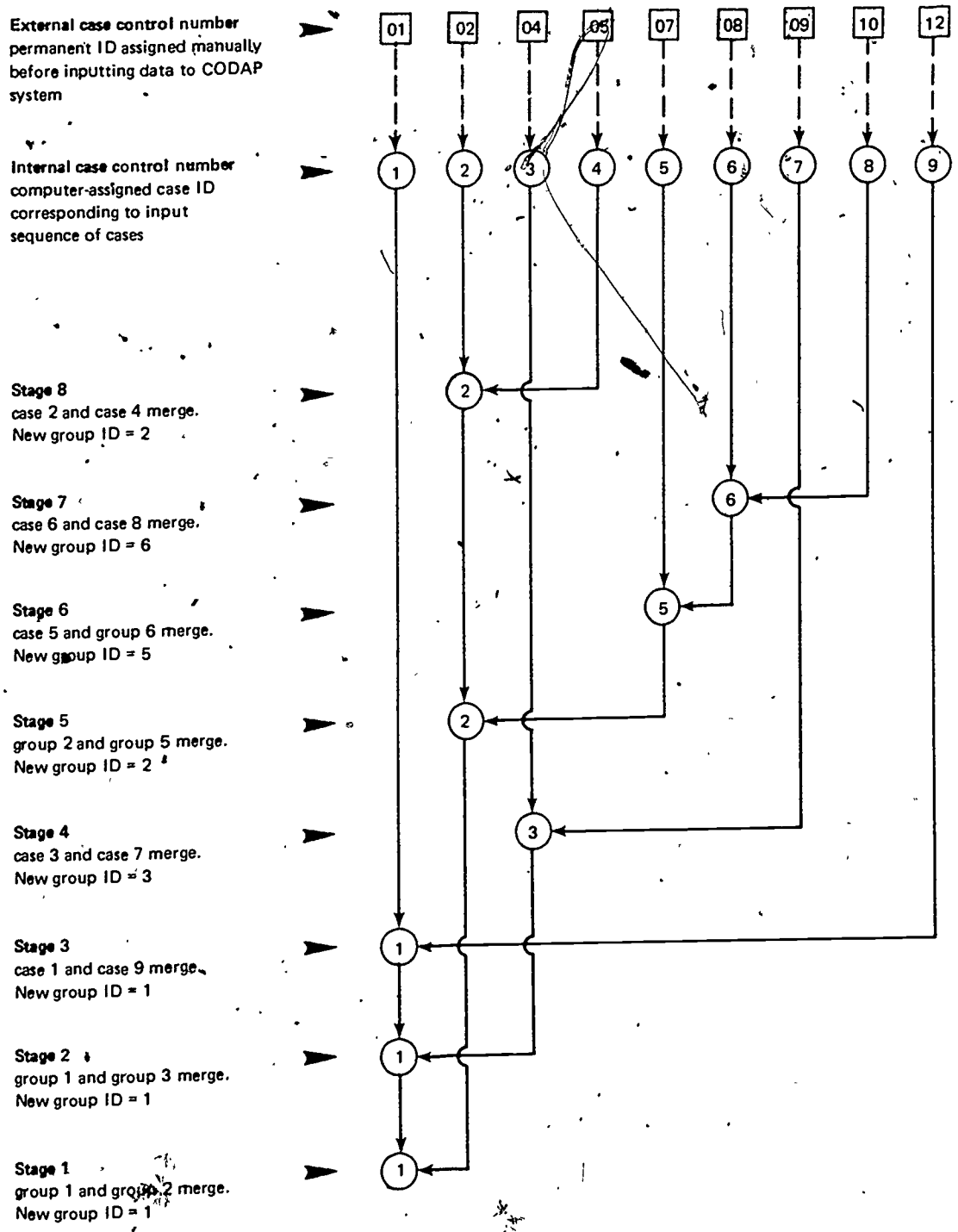

Figure 1. Hierarchical grouping sequence displaying action taken at each stage.

If the nine computer-assigned case numbers (internal case control numbers) are now rearranged so that the grouping process can be displayed as a straight forward diagram, with no branch lines crossing other branch lines, the new sequence is called the "KPATH order" and the sequence numbers assigned to the nine cases are called "KPATH sequence numbers." Figure 2 shows how the branching appears when the nine cases have been reordered in KPATH sequence.

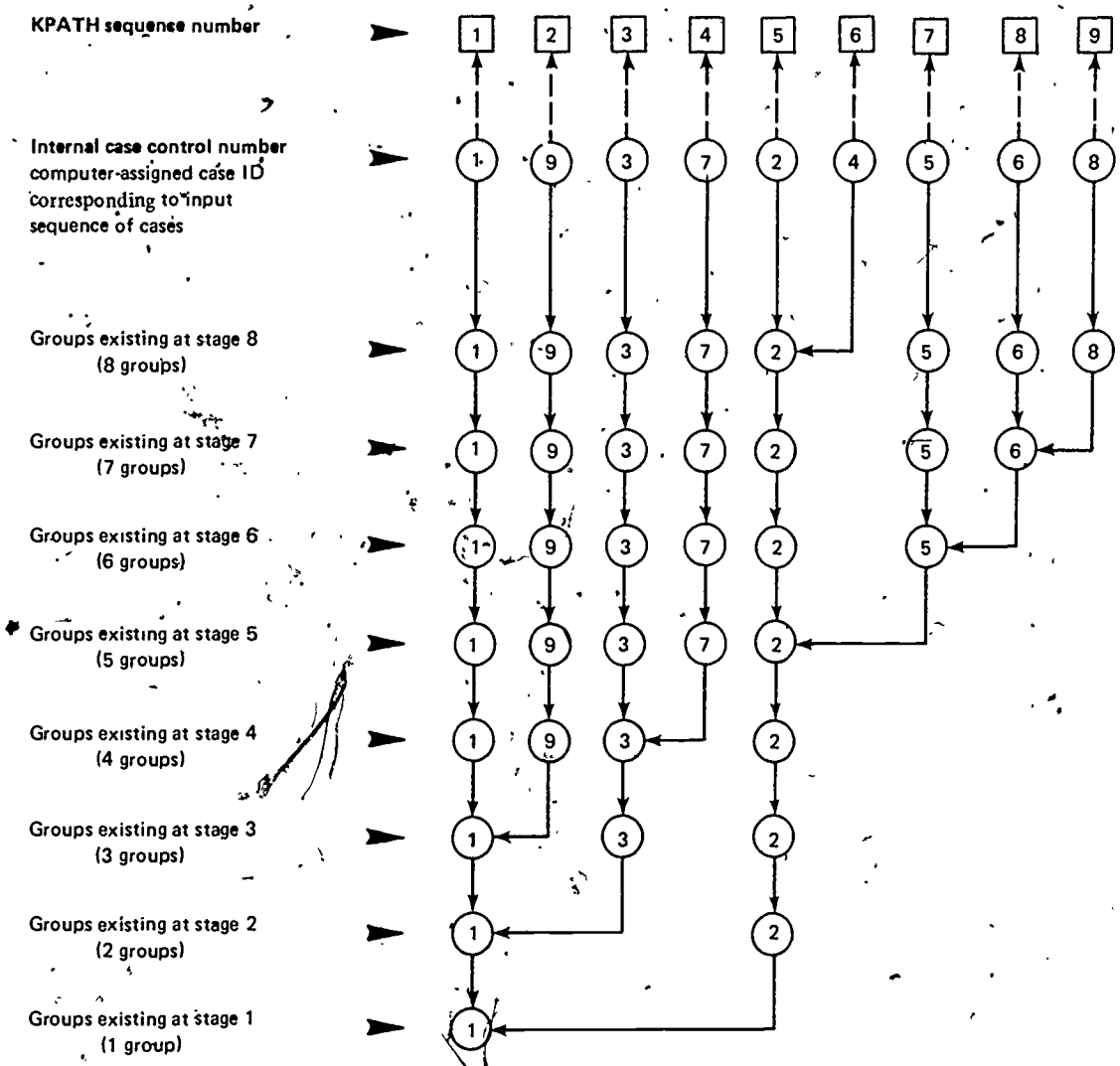

Figure 2. Hierarchical grouping sequence displaying groups existing at each stage.

Figure 3 is a more compressed and visually less complex version of the process shown in Figure 2. It displays *what* groups were formed during the grouping process without indicating the *order* (or stage) at which the groups were formed.

As stated earlier, all groups which merged during any stage of the grouping process will be assigned adjacent KPATH sequence numbers. Inspection of Figure 3 confirms this fact. Cases 1 and 9, for example, were assigned KPATH sequence numbers 1 and 2, respectively, Group 3, which later joined with Group 1, contains cases 3 and 7 and these cases were assigned KPATH sequence numbers 3 and 4, respectively, at that point, Group 1 contained the cases having KPATH sequence numbers 1 through 4, when Groups 1 and 2 merged, Group 2 brought into Group 1 the adjacent block of KPATH sequence numbers, namely, 5 through 9. Thus, the most similar cases and groups have adjoining KPATH sequence numbers. The reordering of case data in KPATH sequence permits identification of background variables having similar

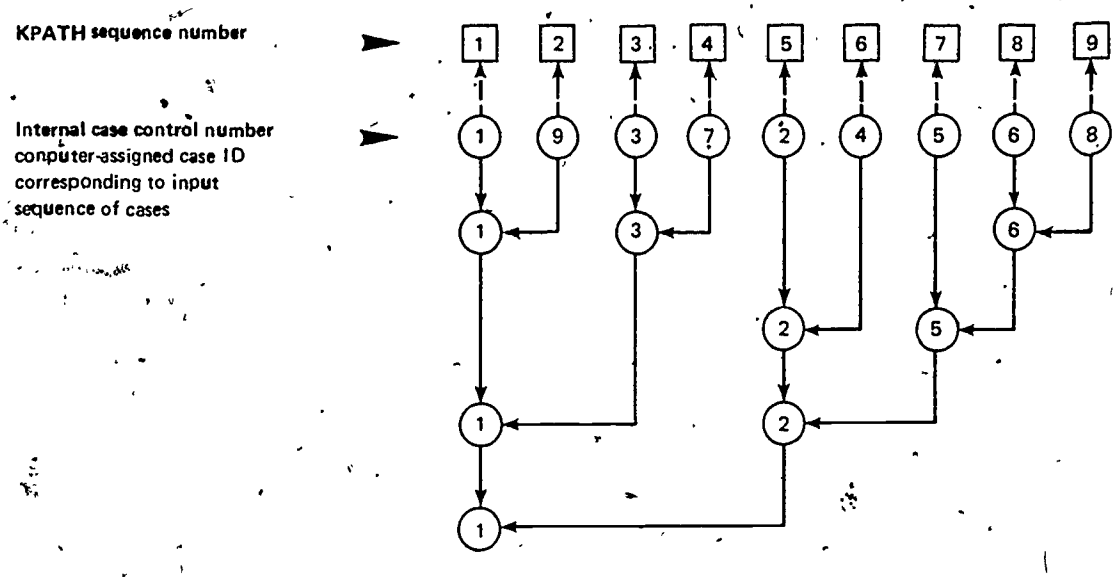

Figure 3. Hierarchical grouping in compressed form.

values for all or most cases in any hierarchical group, and enables the generation of composite job descriptions for hierarchical groups, since the required task data are in a readily accessible sequence.

If the arrowpoints in Figure 3 are viewed as pivots, which allow the circled numbers to flip-flop 180 degrees, it can be seen that there are numerous permutations of groups that would retain the same clustering of individual cases and groups and would yield a valid KPATH sequence. Therefore, the KPATH sequence cannot be viewed as yielding a *continuum* of similarity. No inference can be made concerning the similarity of cases based on their proximity in the KPATH sequence without reference to a complete grouping diagram or a group membership report. Figure 4 illustrates what the KPATH ordering of cases would be if the *highest*-numbered group were placed first throughout the clustering, thus causing the identification numbers for *all* groups to be reversed. The KPATH sequence numbers are now assigned to different case numbers, but the clusters are precisely the same as those in Figure 3.

Up to this point, the clustering process and assignment of KPATH sequence numbers have been presented graphically. The computer, however, does not use diagrams to determine KPATH sequence. It accomplishes this task by developing a series of tables, such as Tables 1, 2, 3, and 4.

The first table the computer produces is one that shows which groups merged at each stage. Comparison of Figure 1 with Table 1 reveals that Groups 2 and 4 did indeed merge at Stage 8; Groups 6 and 8 merged at Stage 7; etc.

The next table to be developed (Table 2) is one which converts the IBEST group numbers to individual (nonrepeated) case numbers. This is done by setting up the IBEST column with values "1" through "9" and reading in the corresponding JBEST values from Table 1. If, however, an IBEST row in Table 2 has been filled with a corresponding JBEST value, and the same IBEST number is encountered again in Table 1, the JBEST value is read into the IBEST row corresponding to the intervening JBEST value. For example, Stage 8 in Table 1 shows IBEST value "2" and JBEST value "4." Table 2 also shows IBEST value "2" and JBEST value "4." But when IBEST value "2" is again encountered at Stage 5 in Table 1, the corresponding JBEST value "5" cannot be inserted next to IBEST value "2" in Table 2, because this cell is already occupied by JBEST value "4." According to the algorithm for forming Table 2, the incumbent JBEST value, i.e., "4," directs the computer to go to IBEST row "4" in Table 2, and, if there is no incumbent JBEST value in IBEST row "4," to insert JBEST value "5" there. Thus, in Table 2, IBEST row "4" shows "5" as the JBEST value. If the JBEST cell in IBEST row "4" had been filled, then that incumbent JBEST value would have directed the computer to another IBEST row for insertion of the unplaced JBEST value. Thus, JBEST value "2" in Table 1 had to be directed to four different IBEST rows in Table 2 before insertion in IBEST row "7." When all JBEST values from Table 1 have been inserted in Table 2, there will be one blank JBEST location remaining, namely, the one in IBEST row "8." A zero is inserted in this JBEST cell. Tables 5 and 6 relate the step-by-step development of Table 2.

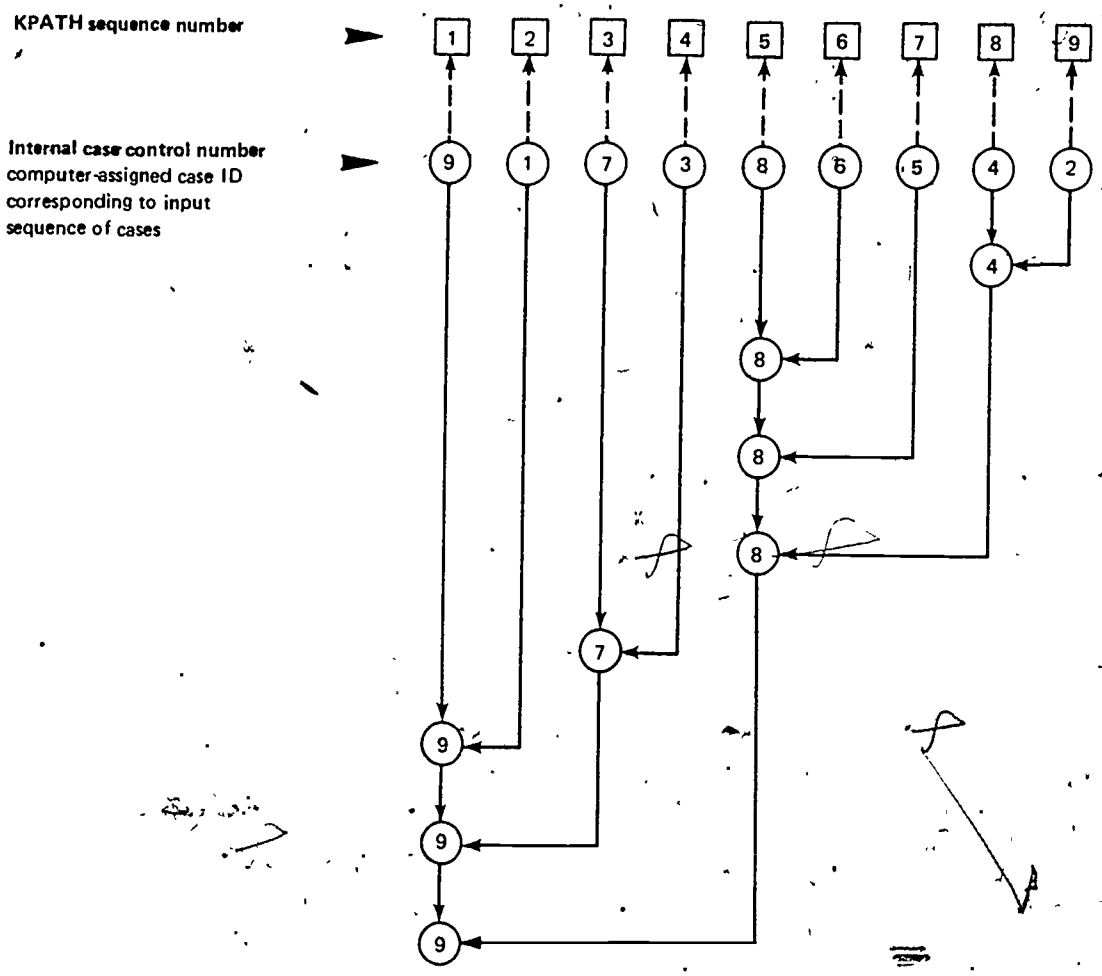

Figure 4. Hierarchical grouping sequence with reversed ordering of groups.

Upon completion of Table 2, Table 3 can be developed from the ordered data. It is in Table 3 that the KPATH sequence number associated with each case ID is generated. This KPATH sequence number specifies the order of the nine cases in Figures 2 and 3.

The procedure for generating Table 3 is as follows. The computer first sets up a column of numbers from "1" to "9" representing the set of KPATH sequence numbers. Then, case number "1" is inserted next to KPATH number "1" in the "Internal Case Control Number" column, because case number "1" is always given KPATH sequence number "1." Next, KPATH sequence number "2" is given to the JBEST value in Table 2 located in IBEST row "1," namely, JBEST value "9." The algorithm for assigning KPATH sequence numbers now directs the computer to go to the IBEST row having the same numeric value as the JBEST value that received the previous KPATH number and assign the next KPATH number to the JBEST value in that IBEST row. Thus, since KPATH number "2" was assigned to JBEST value "9," KPATH number "3" will be assigned to the JBEST value in Table 2 located in IBEST row "9," namely, "3." Moving now to IBEST row "3" in Table 2, the computer selects JBEST value "7" to be assigned KPATH number "4," etc., until each of the nine case numbers has been assigned the appropriate KPATH sequence number. Inspection of Figures 2 and 3 shows that the case number associated with each KPATH sequence number agrees with the association generated in Table 3. Tables 7 and 8 outline the development of Table 3.

Table 4 realigns the internal (computer-assigned) case control number, which is a temporary "group" identifier, with the external or permanent case control number, which, in turn, automatically aligns the external case control number with the appropriate KPATH sequence number. As shown earlier in this report, the correspondence between external and internal case control numbers is established as the time case data is fed into the Input Standard (INPSTD) program.

Table 1. Groups which merged at each stage of the hierarchical grouping

Stage	Groups which merged	
	IBEST	JBEST
8	2	4
7	6	8
6	5	6
5	2	5
4	3	7
3	1	9
2	1	3
1	1	2

Table 2. Conversion of IBEST numbers to case identification numbers

Case identification	
IBEST	JBEST
1	9
2	4
3	7
4	5
5	6
6	8
7	2
8	0
9	3

Table 3. Generation of KPATH sequence numbers corresponding to input case numbers

Internal case control number	KPATH sequence number
1	1
9	2
3	3
7	4
2	5
4	6
5	7
6	8
8	9

Table 4. Alignment of external (permanent case control number) with KPATH sequence number

Case control number		KPATH sequence number
Internal (temporary)	External (permanent)	
1	01	1
9	12	2
3	04	3
4	05	4
2	02	5
4	05	6
5	07	7
6	08	8
8	10	9

Table 5. Logic Diagram for developing Table 2, including tables of input and output data.

Input (Table 1)

Stage	Groups which merged	
	IBEST	JBEST
8	2	4
7	6	8
6	5	6
5	2	5
4	3	7
3	1	9
2	1	3
1	1	2

Output (Table 2)

Case identification	
IBEST	JBEST
1	9
2	4
3	7
4	5
5	6
6	8
7	2
8	0
9	3

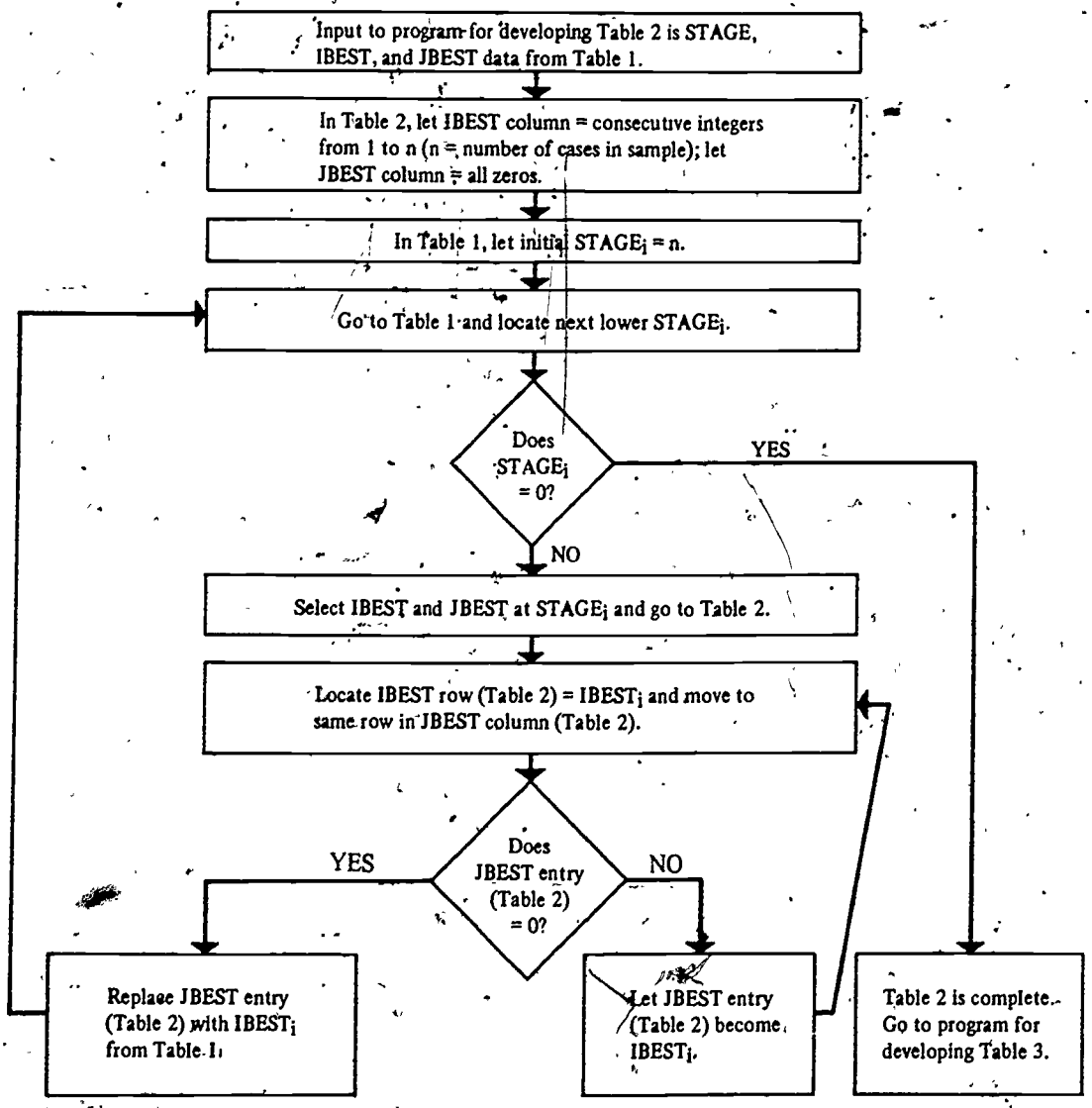

Table 6. Step-by-step development of Table 2 from Table 1 data.

STEP 1. MAKE INITIAL SET-UP OF TABLE 2: IBEST = CONSECUTIVE INTEGERS FROM 1 TO N(n=9); JBEST = ALL ZEROS.

CASE IDENTIFICATION IBEST	JBEST
1.	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0

STEP 2. GO TO TABLE 1, STAGE 8; SELECT IBEST = 2, JBEST = 4; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
1	0
(2)	(4)
3	0
4	0
5	0
6	0
7	0
8	0
9	0

STEP 3. GO TO TABLE 1, STAGE 7; SELECT IBEST = 6, JBEST = 8; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
1	0
2	4
3	0
4	0
5	0
(6)	(8)
7	0
8	0
9	0

STEP 4. GO TO TABLE 1, STAGE 6; SELECT IBEST = 5, JBEST = 6; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
1	0
2	4
3	0
4	0
(5)	(6)
6	8
7	0
8	0
9	0

STEP 5. GO TO TABLE 1, STAGE 5, SELECT IBEST = 2, JBEST = 5; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
1	0
(2)	4
3	0
4	(5)
5	6
6	8
7	0
8	0
9	0

STEP 6. GO TO TABLE 1, STAGE 4, SELECT IBEST = 3, JBEST = 7; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
1	0
2	4
(3)	(7)
4	5
5	6
6	8
7	0
8	0
9	0

STEP 7. GO TO TABLE 1, STAGE 3; SELECT IBEST = 1; JBEST = 9; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
(1)	(9)
2	4
3	7
4	5
5	6
6	8
7	0
8	0
9	0

STEP 8. GO TO TABLE 1, STAGE 2; SELECT IBEST = 1, JBEST = 3; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
(1)	9
2	4
3	7
4	5
5	6
6	8
7	0
8	0
(9)	(3)

STAGE 9. GO TO TABLE 1, STAGE 1; SELECT IBEST = 1, JBEST = 2; INSERT IN TABLE 2.

CASE IDENTIFICATION IBEST	JBEST
(1)	9
2	4
3	7
4	5
5	6
6	8
(7)	(2)
8	0
9	3

Table 7. Logic Diagram for developing Table 3, including tables of input and output data.

Input (Table 2)

Case identification	
IBEST	JBEST
1	9
2	4
3	7
4	5
5	6
6	8
7	2
8	0
9	3

Output (Table 3)

Internal case control number	KPATH sequence number
1	1
9	2
3	3
7	4
2	5
4	6
5	7
6	8
8	9

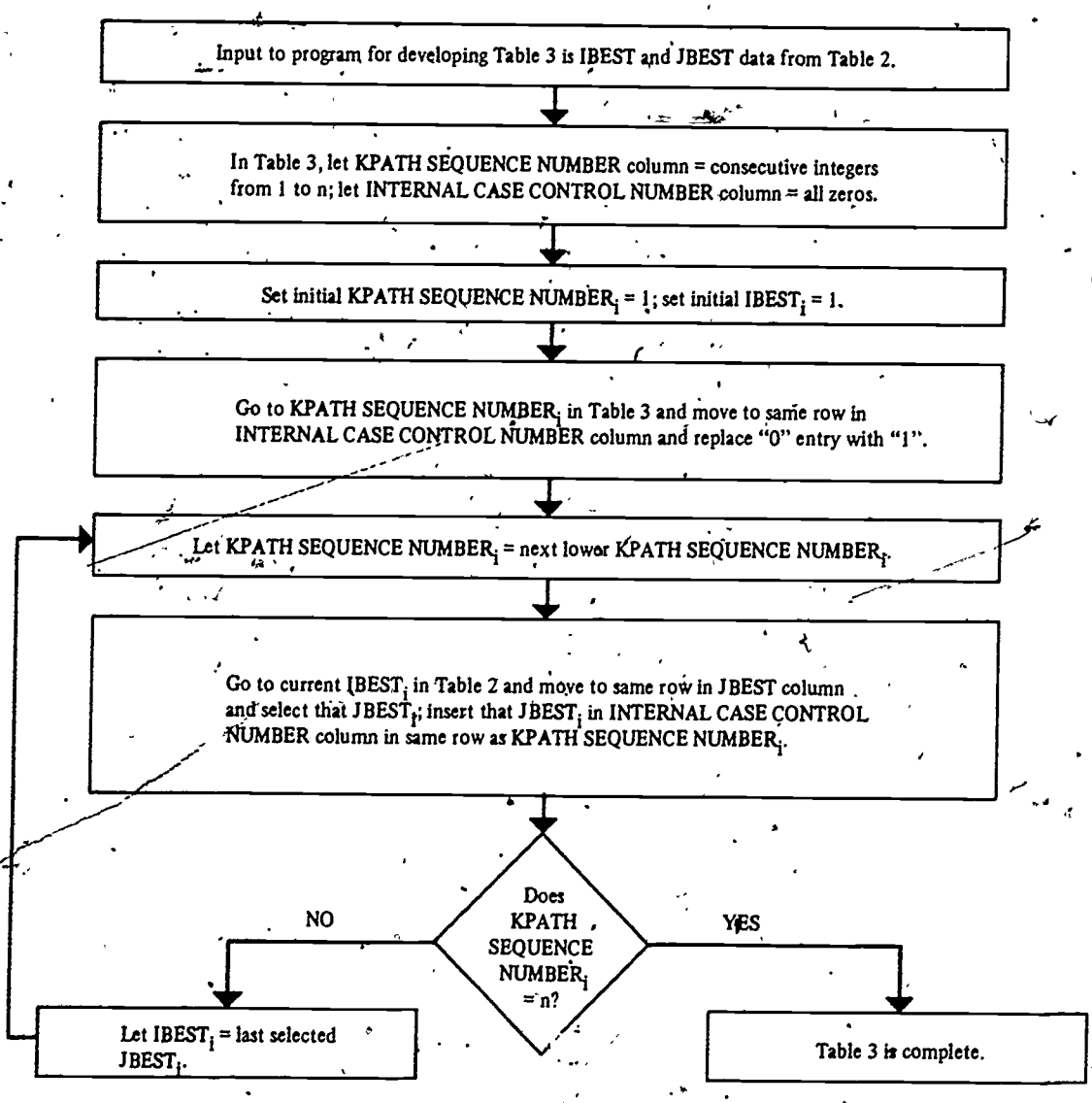

Table 8. Step-by-step development of Table 3 from Table 2 data.

STEP 1. INSERT ZEROS IN INTERNAL CASE CONTROL NUMBER COLUMN OF TABLE 3; INSERT CONSECUTIVE INTEGERS FROM "1" TO "9" IN KPATH SEQUENCE NUMBER COLUMN.

TABLE 2		TABLE 3	
CASE IDENTIFICATION		INTERNAL CASE CONTROL	KPATH SEQUENCE
IBEST	JBEST	NUMBER	NUMBER
1	9	0	1
2	4	0	2
3	7	0	3
4	5	0	4
5	6	0	5
6	8	0	6
7	2	0	7
8	0	0	8
9	3	0	9

STEP 2. INSERT "1" IN INTERNAL CASE CONTROL NUMBER COLUMN OF TABLE 3 OPPOSITE KPATH SEQUENCE NUMBER "1".

TABLE 2		TABLE 3	
CASE IDENTIFICATION		INTERNAL CASE CONTROL	KPATH SEQUENCE
IBEST	JBEST	NUMBER	NUMBER
1	9	1	1
2	4	0	2
3	7	0	3
4	5	0	4
5	6	0	5
6	8	0	6
7	2	0	7
8	0	0	8
9	3	0	9

STEP 3. INTERNAL CASE CONTROL NUMBER "1" INSERTED AT STEP 2 DIRECTS COMPUTER TO GO TO IBEST "1" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "9", WHICH DIRECTS COMPUTER TO INSERT "9" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION		INTERNAL CASE CONTROL	KPATH SEQUENCE
IBEST	JBEST	NUMBER	NUMBER
1	9	1	1
2	4	9	2
3	7	0	3
4	5	0	4
5	6	0	5
6	8	0	6
7	2	0	7
8	0	0	8
9	3	0	9

STEP 4. INTERNAL CASE CONTROL NUMBER "9" INSERTED AT STEP 3 DIRECTS COMPUTER TO GO TO IBEST "9" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "3", WHICH DIRECTS COMPUTER TO INSERT "3" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION		INTERNAL CASE CONTROL	KPATH SEQUENCE
IBEST	JBEST	NUMBER	NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	0	4
5	6	0	5
6	8	0	6
7	2	0	7
8	0	0	8
9	3	0	9

STEP 5. INTERNAL CASE CONTROL NUMBER "3" INSERTED AT STEP 4 DIRECTS COMPUTER TO GO TO IBEST "3" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "7", WHICH DIRECTS COMPUTER TO INSERT "7" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION		INTERNAL CASE CONTROL	KPATH SEQUENCE
IBEST	JBEST	NUMBER	NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	7	4
5	6	0	5
6	8	0	6
7	2	0	7
8	0	0	8
9	3	0	9

STEP 6. INTERNAL CASE CONTROL NUMBER "7" INSERTED AT STEP 5 DIRECTS COMPUTER TO GO TO IBEST "7" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "2", WHICH DIRECTS COMPUTER TO INSERT "2" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION		INTERNAL CASE CONTROL	KPATH SEQUENCE
IBEST	JBEST	NUMBER	NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	7	4
5	6	2	5
6	8	0	6
7	2	0	7
8	0	0	8
9	3	0	9

Table 8. Step-by-step development of Table 3 from Table 2 data, (Cont'd).

STEP 7. INTERNAL CASE CONTROL NUMBER "2" INSERTED AT STEP 6 DIRECTS COMPUTER TO GO TO IBEST "2" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "4", WHICH DIRECTS COMPUTER TO INSERT "4" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION IBEST	JBEST	INTERNAL CASE CONTROL NUMBER	KPATH SEQUENCE NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	7	4
5	6	2	5
6	8	4	6
7	2	0	7
8	0	0	8
9	3	0	9

STEP 8. INTERNAL CASE CONTROL NUMBER "4" INSERTED AT STEP 7 DIRECTS COMPUTER TO GO TO IBEST "4" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "5", WHICH DIRECTS COMPUTER TO INSERT "5" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION IBEST	JBEST	INTERNAL CASE CONTROL NUMBER	KPATH SEQUENCE NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	7	4
5	6	2	5
6	8	4	6
7	2	5	7
8	0	0	8
9	3	0	9

STEP 9. INTERNAL CASE CONTROL NUMBER "5" INSERTED AT STEP 8 DIRECTS COMPUTER TO GO TO IBEST "5" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "6", WHICH DIRECTS COMPUTER TO INSERT "6" AS THE NEXT INTERNAL CASE CONTROL NUMBER.

TABLE 2		TABLE 3	
CASE IDENTIFICATION IBEST	JBEST	INTERNAL CASE CONTROL NUMBER	KPATH SEQUENCE NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	7	4
5	6	2	5
6	8	4	6
7	2	5	7
8	0	6	8
9	3	0	9

STEP 10. INTERNAL CASE CONTROL NUMBER "6" INSERTED AT STEP 9 DIRECTS COMPUTER TO GO TO IBEST "6" IN TABLE 2, THENCE TO THE CORRESPONDING JBEST "8", WHICH DIRECTS COMPUTER TO INSERT "8" AS THE NEXT INTERNAL CASE CONTROL NUMBER. TABLE 3 IS NOW COMPLETE.

TABLE 2		TABLE 3	
CASE IDENTIFICATION IBEST	JBEST	INTERNAL CASE CONTROL NUMBER	KPATH SEQUENCE NUMBER
1	9	1	1
2	4	9	2
3	7	3	3
4	5	7	4
5	6	2	5
6	8	4	6
7	2	5	7
8	0	6	8
9	3	8	9

PRKPTH Program

The Print KPATH (PRKPTH) program[3] produces a report of case data values for selected background and computed variables in case ID or KPATH order. PRKPTH obtains the required data from either the history data tape or the KPATH tape, depending on the desired sequence of the reported data. The PRKPTH report in KPATH order is especially useful in identifying similarity of background characteristics for groups formed by the hierarchical clustering process. This capability of PRKPTH aids the job analyst in discerning which groups on the cluster diagram printout (DIAGRM) are most susceptible to definition as meaningful job types. The KPATH ordering of background data is also useful in understanding and explaining the worker composition of a selected job-type cluster. The following discussion and the accompanying excerpts from various PRKPTH reports explain and illustrate how PRKPTH is used to locate and define job-type groups. Table 7 is a portion of a PRKPTH report displaying KPATH-ordered background data from a Security Police (811X0) Airman Job Inventory. The first two columns of every PRKPTH report list the "KPATH Sequence Number" and the "Case Control Number," respectively. The KPATH sequence number automatically appears in every PRKPTH report and the case control number appears as a matter of policy.

Several of the columns in Table 9 provide insight into the nature of the hierarchical groups that were formed on the basis of similarity of time spent on tasks listed in the job inventory. For example, column 3, "Major Command," reveals that there is a job cluster peculiar to SAC, designated as "S," running from KPATH number 1315 to KPATH number 1318 and from KPATH number 1324 to KPATH number 1331. There appears to be another command-related job cluster at KPATH number 1333 through KPATH number 1350, the code letter "R" indicating PACAF.

It appears that there are two very distinct clusters of cases, KPATH 1301–1320 and KPATH 1321–1350, that may be delineated on several variables. The cluster running from KPATH 1301–1320 is generally higher in grade (Column 4), duty AFSC (Column 6), number of months of active Federal military service (Column 12), and number of subordinates directly supervised (Column 13). This cluster also has a subcluster (KPATH 1302–1308) containing supervisors who are mainly in DAFSC 81170. On the other hand, every member of the KPATH 1321–1350 cluster is military dog qualified, as indicated by the "A" suffix attached to the duty AFSC.

Table 10 is a portion of a PRKPTH report which displays categorical responses of KPATH-order job incumbents in the Supply Services (61XXX) Airman Career Ladder. The "1" response designates those "areas" (broad work activities) in which an incumbent spends the most time in his present assignment. There are several patterns of "1's" that indicate probable job-type clusters. One group identified by KPATH numbers 251 to 263 works in a variety of areas in the commissary. Most of the incumbents in this group work in at least two commissary work areas. Another group (KPATH 265–273) is confined to the commissary warehouse. A third group (KPATH 279–298) works primarily at a remote or isolated Base Exchange.

Table 11 is a portion of a PRKPTH report that displays job titles descriptive of incumbents' present work assignments. These KPATH-ordered titles, which are descriptive of jobs in the Inventory Management and Materiel Facilities (64XX0) Airman Career Ladders, were furnished by the surveyed incumbents. Inspection of job titles for communality of labeling revealed four clearcut job-type clusters: KPATH 304–312 is a group of NCOIC's of a demand processing unit; KPATH 313–323, a group of NCOIC's of a materiel control section; and KPATH 339–350, a group of NCOIC's of a requirements and requisitioning unit. Of course, not all the titles within each group correspond to the general title attributed to the group, but there is sufficient agreement for one to assume that deviant titles indicate nonstandard or inaccurate labeling.

Tables 9, 10, and 11 provide only a small sampling of the various kinds of background data that may appear on a PRKPTH report. Other PRKPTH's might include equipment used, technical manuals used, technical courses completed, highest level of education, job satisfaction ratings, reenlistment intent, job utilization of talents and training, past experience in various work areas, organizational level of position, functional areas requiring training, method by which assigned to present career ladder, previous AFSC of cross-trainee, assigned organization, job location, etc.

When data such as these are KPATH-ordered, not only are static characteristics of the hierarchically formed job clusters revealed, but also something of the interaction dynamics of relevant job variables within diverse job types. This type of information has meaningful application to many job-related problems, among which are the development and revision of technical training courses, the interface of job assignments with career progression objectives, the reengineering of jobs to meet specific needs, and the pinpointing of sources of job satisfaction or dissatisfaction.

III. SUMMARY

This paper has described two CODAP programs: the Ordering of Hierarchically Grouped Case Data (KPATH) program and the Print KPATH (PRKPTH) program. In each instance, examples have been given to assist the occupational analyst in learning how these programs may serve his needs.

Table 9. Portion of a KPATH printout displaying selected background information for hierarchically ordered cases from the Security Police (811X0) Airman Job Inventory.

KPATH SEQ NUM	CASE CTRL NUM	MAJ CMD	GRD	PRIMARY AFSC	DUTY AFSC	NUM TASKS PERFM	MOS DUTY AFSC	MOS PRES BASE	MOS PRES ASGMT	MOS CAREER FIELD	MOS ACT FED MIL SERVICE	NUMBER SUPER-VISED
1301	4038	R	4	81150	81150	14	162	030	162	162	162	++
1302	0935	S	5	81170	81170	37	024	016	010	230	241	03
1303	4082	A	5	81170	81170	50	072	009	002	190	192	05
1304	1139	C	5	81170	81170	21	+++	038	017	186	189	03
1305	4166	S	5	81170	81150	24	110	011	003	118	121	04
1306	2142	R	6	81170	81170	27	030	009	008	237	241	04
1307	3825	A	4	81150	+++++++	26	028	023	005	028	028	++
1308	2569	R	4	81150A	81170	32	001	002	001	071	071	06
1309	4701	S	4	81150	81150	55	051	007	004	069	069	++
1310	0669	S	5	81170	81150	16	002	002	002	112	114	04
1311	2829	D	3	81130	81150	26	008	004	003	008	010	++
1312	4356	R	5	+++++++	+++++++	15	142	004	004	147	235	++
1313	4478	+	3	81130	81150	10	016	015	014	016	020	++
1314	2545	R	5	81170	81170	13	006	003	001	185	190	05
1315	4124	S	4	81150	81150	23	040	003	002	042	043	02
1316	4418	S	+	81150	81150	20	011	003	027	026	027	++
1317	4731	S	3	81150	81150	34	024	002	+++	034	034	++
1318	0938	S	5	81150	81150	23	120	010	006	120	150	++
1319	1438	R	6	81170	81170	36	104	004	004	207	207	05
1320	1542	R	4	81150	81150	43	022	018	006	034	035	++
1321	0299	C	4	81150A	81150A	47	036	026	007	042	043	++
1322	3666	S	3	81130A	81130A	52	007	007	007	010	012	++
1323	3683	+	3	81130A	81150A	43	005	007	007	007	010	++
1324	0726	S	3	81150A	81150A	38	014	014	001	014	016	++
1325	0737	S	3	81130A	81130A	36	012	012	007	012	014	++

Table 9. Portion of a KPATH printout displaying selected background information for hierarchically ordered cases from the Security Police (811X0) Airman Job Inventory. (Cont'd)

KPATH SEQ NUM	CASE CTRL NUM	MAJ CMD	GRD	PRIMARY AFSC	DUTY AFSC	NUM TASKS PERFM	MOS DUTY AFSC	MOS PRES BASE	MOS PRES ASGMT	MOS CAREER FIELD	MOS ACT FED MIL SERVICE	NUMBER SUPER-VISED
1326	4649	S	3	81150A	81150A	46	021	028	021	027	028	++
1327	0735	S	3	81150A	81150A	40	014	014	012	014	050	++
1328	1663	S	4	81150A	81150A	38	011	018	018	045	045	++
1329	0727	S	4	81150A	81150A	39	007	025	007	025	026	++
1330	2463	S	3	81130A	81150A	32	018	018	018	018	020	++
1331	4646	S	3	81130A	81150A	35	030	007	007	036	036	++
1332	0783	+	3	81130A	81150A	41	007	007	007	014	014	++
1333	1650	R	3	81130A	81130A	41	014	006	014	014	018	++
1334	0786	R	4	81150A	81150A	37	024	006	006	036	036	++
1335	1401	S	3	81150A	81150A	35	034	019	+++	039	042	++
1336	2514	+	2	81130A	81150A	43	005	005	002	005	007	++
1337	0799	R	+	81150A	81150A	38	017	006	006	030	035	++
1338	0877	R	3	81150A	81150A	37	017	009	009	020	022	++
1339	1720	R	4	81150A	81150A	36	015	009	009	032	032	++
1340	3530	+	3	81150A	81150A	40	015	001	001	024	025	++
1341	2124	R	3	81150A	81150A	43	015	007	007	020	024	++
1342	4695	S	3	81150A	81150A	40	007	007	007	009	013	++
1343	3667	S	3	81130A	81130A	41	006	007	006	006	008	++
1344	0878	R	3	81150A	81150A	43	032	009	032	045	045	++
1345	0879	R	3	81150A	81150A	45	015	009	015	026	026	++
1346	1716	R	4	81150A	81150A	46	045	009	032	046	046	++
1347	1696	R	3	81150A	81150A	43	022	007	006	022	023	++
1348	1698	R	4	81150A	81150A	43	034	002	002	042	042	02
1349	1697	+	3	81150A	81130A	42	022	008	008	022	023	++
1350	1700	R	4	81150A	81150A	51	029	007	024	029	032	02

Table 10. Portion of a KPATH printout displaying selected work area information for hierarchiacally ordered cases from the supply services (61XXX) Airman Job Inventory.

KPATH SEQ NUM	CASE CTRL NUM	******AREAS IN WHICH MOST TIME SPENT-PRESENT ASSIGNMENT****** STAFF LEVEL	***EXCHANGE*** CONUS	O/S	SITE	HOUS ING	HOUSG SUPP	*******COMMISSARY******* ANNEX	SALES	STKNG	WREHSE
251	0025	0	0	0	0	0	0	0	1	0	0
252	0228	0	0	0	1	0	0	0	1	1	1
253	0229	0	0	0	1	0	0	0	1	1	0
254	0804	0	0	0	0	0	0	0	0	1	0
255	0702	0	0	0	0	0	0	0	1	1	1
256	1203	0	0	0	0	0	0	1	1	1	1
257	0821	0	0	0	0	0	0	0	0	0	0
258	0377	0	0	0	1	0	0	0	1	1	0
259	0852	0	0	0	0	0	0	0	0	1	1
260	0996	0	0	0	0	0	0	0	1	0	0
261	1195	0	0	0	0	0	0	0	1	1	0
262	0329	0	0	0	0	0	0	1	0	1	0
263	1000	0	0	1	0	0	0	0	1	1	1
264	1280	0	0	1	0	0	0	0	0	0	0
265	0250	0	0	0	0	0	0	0	0	0	1
266	0758	0	0	0	0	0	0	0	0	0	1
267	0820	0	0	0	0	0	0	0	0	0	1
268	0841	0	0	0	0	0	0	0	0	0	1
269	0080	0	0	0	0	0	0	0	0	0	1
270	0095	0	0	0	0	0	0	0	0	0	1
271	0701	0	0	0	0	0	0	0	0	0	1
272	0868	0	0	0	0	0	0	0	0	0	1
273	0218	0	0	0	0	0	0	0	0	0	1
274	1016	0	0	0	0	0	0	0	1	1	0
275	0392	0	0	0	0	0	0	0	1	0	0

Table 10. Portion of a KPATH printout displaying selected work area information for hierarchically ordered cases from the supply services (61XXX) Airman Job Inventory. (Cont'd)

KPATH SEQ NUM	CASE CTRL NUM	******AREAS IN WHICH MOST TIME SPENT-PRESENT ASSIGNMENT****** STAFF LEVEL	***EXCHANGE*** CONUS	O/S	SITE	HOUS ING	HOUSG SUPP	*******COMMISSARY******* ANNEX	SALES	STKNG	WREHSE
276	0498	0	0	0	0	0	0	0	1	0	0
277	0708	0	0	0	0	0	0	0	0	1	0
278	0735	0	0	0	0	0	0	0	0	1	0
279	0008	0	0	0	1	0	0	0	0	0	0
280	0057	0	0	0	1	0	0	0	1	0	0
281	0038	0	0	0	1	0	0	0	0	0	0
282	0970	0	1	0	1	0	0	0	0	0	0
283	0326	0	0	0	1	0	0	0	0	0	0
284	0129	0	0	0	1	0	0	0	0	0	0
285	0234	0	0	0	1	0	0	0	0	0	0
286	0973	0	0	0	1	0	0	0	0	0	0
287	0764	0	0	0	1	0	0	0	1	0	0
288	0931	0	0	0	0	0	0	1	0	0	0
289	0009	0	0	0	1	0	0	0	0	0	0
290	0124	0	0	0	1	0	0	0	0	0	0
291	0822	0	0	0	1	0	0	0	0	0	0
292	0859	0	0	0	1	0	0	0	0	0	0
293	0036	0	0	0	1	0	0	0	0	0	0
294	0093	0	0	0	1	0	0	0	0	0	0
295	1202	0	0	0	1	0	0	0	0	1	1
296	0760	0	0	0	1	0	0	0	0	0	0
297	1208	0	0	0	1	0	0	0	0	0	0
298	0048	0	0	0	1	0	0	0	1	1	0
299	0429	0	0	0	0	0	0	0	1	0	0
300	0309	0	0	0	0	0	0	0	1	1	0

Table 11. Portion of a KPATH printout displaying present work assignment job titles of hierarchically ordered cases from the inventory management/materiel facilities (64XX0) Airman Job Inventory.

KPATH SEQ NUM	CASE CTRL NUM	PRESENT WORK ASSIGNMENT
301	1794	NCOIC REPAIR CYCLE SUPPORT
302	1478	DIFM CONTROL
303	1667	NCOIC WORK ORDERS AND ISSUE POINT MAINT SUPPORT
304	0023	NCOIC DEMAND PROCESSING INVENTORY MANAGEMENT SUPERVISOR
305	0510	INVT MGMT SUPERVISOR NCOIC DEMAND PROCESSING
306	1686	DEMAND PROCESSING SHIFT SUPERVISOR
307	1826	DEMAND PROCESSING SUPERVISOR EXPEDITE
308	0348	NCOIC MAINTENANCE SUPPORT SECTION
309	0903	NCOIC DEMAND PROCESSING
310	1259	CHIEF OF DEMAND PROCESSING SMO
311	1797	NCOIC SUPPLIES DEMAND PROCESSING UNIT
312	1571	NCOIC DEMAND PROCESSING UNIT
313	0132	MATERIEL CONTROL SUPERVISOR
314	1658	NCOIC MATERIEL CONTROL
315	0495	REQUIREMENTS MATERIEL CONTROL INVENTORY MANAGEMENT SUPERVISOR
316	0752	NORS MONITOR MATERIAL CONTROL
317	0500	NCOIC SUPPORT RECORDS
318	1886	NCOIC REQUIREMENTS SECTION
319	0353	
320	0719	NCOIC FORWARD SUPPLY SUPPORT ENGINE PROPELLER DIVISION
321	1059	NCOIC EQUIPMENT CONTROL
322	0320	NCOIC MATERIEL CONTROL ACTIVITY
323	0342	CHIEF MATERIEL CONTROL BASE MAINTENANCE
324	0164	
325	2158	NCOIC MAINT SUPPORT SUPPLIES MGT DIVISION CHIEF OF SUPPLY

Table 11. Portion of a KPATH printout displaying present work assignment job titles of hierarchically ordered cases from the inventory management/materiel facilities (64XX0) Airman Job Inventory. (Cont'd)

KPATH SEQ NUM	CASE CTRL NUM	PRESENT WORK ASSIGNMENT
326	1066	PRIORITY MONITOR
327	1651	
328	1652	NCOIC BENCH STOCK TCTO SECTION
329	1656	NORS CONTROL ENGINE MONITOR
330	0026	NCOIC MATERIEL CONTROL CIVIL ENGINEER
331	0125	NCOIC SUPPLIES MGT BRANCH
332	2224	NCOIC STOCK CONTROL SECTION
333	0127	NCOIC STOCK CONTROL SECTION SUPPLIES MANAGEMENT BRANCH
334	0350	STOCK CONTROL SUPERVISOR
335	1128	NCOIC REQUIREMENTS UNIT
336	2127	NCOIC STOCK CONTROL
337	1895	NCOIC STOCK CONTROL
338	0396	NCOIC EMO
339	0063	NCOIC OF EMO REQUIREMENT AND REQUISITION UNIT
340	0356	REQUIREMENTS AND REQUISITION BRANCH NCOIC
341	1011	NCOIC REQUIREMENT AND REQUISITION EMO
342	1966	REQUIREMENTS AND REQUISITIONING
343	1561	ASST NCOIC REQUIREMENTS AND REQUISITIONING OF EMO
344	0380	
345	0499	NCOIC REQUIREMENT AND REQUISITIONING UNIT (EMO)
346	2229	NCOIC REQUIREMENTS UNIT STOCK CONTROL
347	1906	
348	0939	REQUISITIONING CLERK
349	0062	NCOIC NORS CONTROL SECTION
350	1718	NCOIC REQUISITION UNIT

REFERENCES

AF Manual 35-2. *Occupational analysis Procedures for surveying and evaluating Air Force Specialities.* Washington: Department of the Air Force, 9 July 1968.

Archer, W. B., & Fruchter, D. A. *The construction, review, and administration of Air Force job inventories.* PRL-TDR-63-21. AD-426 755. Lackland AFB, Tex.. Personnel Research Laboratory, Aerospace Medical Division, August 1963.

Archer, W. B. *Computation of group job descriptions from occupational survey data.* PRL-TR-66-12, AD-653 543. Lackland AFB, Tex.. Personnel Research Laboratory, Aerospace Medical Division, December 1966.

Christal, R. E., & Ward, J. H., Jr. The MAXOF clustering model. In M. Lorr & S. B. Lyerly (Eds.), *Proceedings of the conference on cluster analysis of multivariate data.* New Orleans. Catholic University of America, June 1967. 11.02-11.45.

Christal, R. E. *CODAP input standard (INPSTD) and variable generation (VARGEN) programs.* AFHRL-TR-72-51. AD-750 144. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, May 1972..

Computer Sciences Corporation. *Program description manual - CODAP Computerized occupational data analysis programs.* (5th ed.) El Segundo, Calif.. Computer Sciences Corporation, September 1969.

Computer Sciences Corporation. *Program description manual DIAGRM - Computerized diagram of hierarchical grouping actions.* (supplement to *Program description manual - CODAP Computerized occupational data analysis programs* (5th ed.)). El Segundo, Calif.. Computer Sciences Corporation, September 1969.

Morsh, J. E. *Identification of job types in the personnel career field.* PRL-TR-65-9, AD-622 433. Lackland AFB, Tex.: Personnel Research Laboratory, Aerospace Medical Division. April 1965.

Morsh, J. E., & Christal, R. E. *Impact of the computer on job analysis in the United States Air Force.* PRL-TR-66-19, AD-656 304. Lackland AFB, Tex.. Personnel Research Laboratory, Aerospace Medical Division. October 1966.

Morsh, J. E., & Archer, W. B. *Procedural guide for conducting occupational surveys in the United States Air Force.* PRL-TR-67-11, AD-664 036. Lackland AFB, Tex.. Personnel Research Laboratory, Aerospace Medical Division, September 1967.

Phalen, W. J., & Christal, R. E. *Comprehensive occupational data analysis programs (CODAP), group membership (GRMBRS/GRPMBR) and automated diagramming (DIAGRM) programs.* AFHRL-TR-73-5, AD-767 199. Lackland AFB, Tex.. Personnel Research Division, Air Force Human Resources Laboratory, April 1973.

Ward, J. H., Jr. *Hierarchical grouping to maximize payoff.* WADD-TN-61-29, AD-261 750. Lackland AFB. Tex.: Personnel Laboratory, Wright Air Development Division, March 1961.

FOOTNOTES

1. The name "KPATH" is derived as follows. The letter "K" represents an integer, and the term "PATH" denotes a queuing operation. The KPATH program uses a series of integers to identify groups being formed during the hierarchical clustering process, such that a path of integers is established which identifies the membership of any previously formed group at a subsequent stage of the clustering process.

2. The clustering action occurring at each stage of GROUP is displayed in reports generated by the Group Membership (GRMBRS) and Automated Diagramming (DIAGRM) programs, which are explained in Phalen and Christal (1973). Also see definitions of "similarity matrix" and "hierarchical grouping" in Appendix B.

3. After this report had been completed in its present form, the PRKPTH program was revised to include the following options:

   a. printing only that portion of the KPATH data that pertains to cases in a specified job group;

   b. reordering KPATH data from high-to-low or low-to-high on one or two selected variables. As a consequence, the title "Print KPATH" (PRKPTH) was changed to "Print Variables" (PRTVAR).

APPENDIX A:
DESCRIPTION OF SELECTED CODAP PROGRAMS REFERRED TO IN THIS REPORT

CODAP: COMPREHENSIVE OCCUPATIONAL DATA ANALYSIS PROGRAMS

CODAP is a computerized occupational data analysis system which inputs and performs calculations upon raw data from job inventories. It is designed to furnish users with a wide variety of reports that facilitate the identification of individual and group job characteristics and the detection of between-job similarities and differences.

INPSTD: RAW DATA EDITING AND INPUT

This program reads task titles, task responses, and background data from tape or card input. It edits the data, converts the raw task responses to percentages, constructs data vectors for each case, reorganizes the data to a standard history data format, and writes the formatted data on the output tape for use in subsequent programs. INPSTD will accept a maximum of: 1700 background and/or computed variables, 1700 task variables, 26 duty variables, and 20,000 cases.

OVRLAP: RELATING RESPONSES TO EACH OTHER

This program generates an overlap or similarity matrix of all possible paired comparisons between individual cases. Similarity is expressed as a percentage of common tasks performed (TSKOVL) or as the total overlapping percentage of time spent on tasks (TIMOVL). Overlap in terms of percent time spent is the preferred option in most studies. OVRLAP can handle up to 3,200 cases and 1,700 tasks.

KPATH: ORDERING A HISTORY DATA TAPE TO REFLECT OUTCOME OF HIERARCHICAL GROUPING

After OVRLAP and GROUP have been completed, the KPATH program assigns sequence numbers to individual cases in such a way that each pair of individuals or groups merged during the grouping process will have a contiguous block of KPATH sequence numbers.

PRKPTH: PRINTING A HISTORY DATA REPORT[3]

PRKPTH enables the user to select variables and printout formats to produce a report of the case data values for selected background and computed variables. The data to be printed is obtained from a history data tape which may be in case ID or KPATH order. The data is not sorted and, therefore, the output will be in case ID or KPATH order.

GRMBRS: REPORTING GROUP MEMBERSHIP

This program produces a report that identifies the two groups combining at each stage of the hierarchical grouping process. The information includes: stage number, number of members in the combined group, number of members in the combining groups, range of KPATH sequence numbers for the combining groups, average percentage of overlap between members of the combining groups, and the average percentage of overlap within the combined group.

DIAGRM: GRAPHICAL PRESENTATION OF HIERARCHICAL GROUPING ACTIONS

This program uses data from the GRMBRS program to generate a tree-like diagram that visually displays the order in which groups merged during the hierarchical grouping process. Each group is

represented by a rectangular block of data containing most of the information reported in GRMBRS. Rows and columns of asterisks show the branches leading from a group to its subgroups. Control values are used to limit the number and type of groups displayed by DIAGRM.

JOBGRP/JOBSPC: CALCULATING COMPOSITE JOB DESCRIPTIONS

This program calculates and prints composite job descriptions for groups formed during the hierarchical grouping process (JOBGRP) or for special groups whose membership is defined in terms of background or computed variables (JOBSPC). Both duty and task job descriptions may be reported in high to low sequence of either "average percent time spent by all members" or "percent of members performing." (A duty is a designated functional area comprising a number of tasks.)

A job description produced by JOBGRP or JOBSPC provides the following information: duty/task number, duty/task title, percent of members performing each duty/task, average percent time spent by members performing each duty/task, average percent time spent on each duty/task by all members, and cumulative average percent time spent by all members.

APPENDIX B:
DEFINITION OF TERMS USED IN THE KPATH PROGRAM DOCUMENTATION

External Case Control Number. A numerical ID manually assigned to each case in a sample such that no two cases have the same ID.

Hierarchical Grouping. Grouping (or clustering) procedures are designed to group large numbers of persons, objects, jobs, etc., into smaller numbers of mutually exclusive classes in which the members have similar characteristics. When the grouping establishes a taxonomy of mutually exclusive clusters wherein each larger unit is a unique combination of the next-subordinate units, the clusters are called "hierarchical groups."

IBEST. The group to be merged at each stage of the clustering process which has the *lower* group ID number.

Internal Case Control Number. A sequential numerical ID assigned by the computer to each case during input. (Corresponds to external case control number if none of the sample cases has been previously removed.)

Internal Sequence Number. The "last stage encountered" (cf. below) plus one.

JBEST. The group to be merged at each stage of the clustering process which has the *higher* ID number.

KPATH Sequence Number. The numerical ID assigned by the computer to each case after the clustering process has been completed such that those cases (or groups of cases) which merged at each stage of the hierarchical clustering process are assigned adjacent numbers.

Last Stage Encountered. The stage at which an IBEST group loses its ID.

Similarity Matrix. An n-case X n-case symetrical matrix containing similarity indices expressed as percentages of overlap between all possible pairs of cases. The percentage overlap values may be in terms of time spent or tasks performed. After the first grouping stage (when the most similar pair of cases has been merged into a single group through an averaging process), the matrix is collapsed into an (n-1) X (n-1) case matrix. The grouping and collapsing process continues until all cases have been merged into a single group represented by a 1 X 1 matrix.

Stage or Stage Number. The point in the clustering process at which a group is formed by the merging of two mutually exclusive subordinate groups.

☆U.S. GOVERNMENT PRINTING OFFICE: 1975-671-602/8004